

EXHIBIT 2

From: [Edward A. Bayley](#)
To: [Reetu Sinha](#); [Alejandra Salinas](#); [openaicopyrightlitigation.lwteam@lw.com](#); [KVP-OAI](#)
Cc: [Alter-TT \(Alter-AI-TT@simplelists.susmangodfrey.com\)](#); [OpenAICopyright](#)
Subject: RE: AG/OAI: Source Code
Date: Sunday, January 12, 2025 11:43:20 PM

Reetu,

As I have explained several times now on this email chain, OpenAI is willing to agree to make available for inspection source code that is responsive to a valid request for production served pursuant to the Rules of Civil Procedure. The bullet points in Alejandra's email below, however, are not RFPs. On Thursday's call, Plaintiffs identified—for the very first time—fourteen requests for production served by Plaintiffs that that it now contends encompass the source code described in Alejandra's bullets—RFPs that could have been identified earlier. As I indicated on the call, I will endeavor to respond Plaintiffs' contention about the scope of these newly identified RFPs by next week, but making a demand after our call for an immediate response over the weekend is patently unreasonable.

We will respond to Plaintiffs' new set of RFPs in due course.

Regards,

-Ed

Edward A. Bayley
Keker, Van Nest & Peters LLP
633 Battery Street
San Francisco, CA 94111-1809
415 773 6672 direct | 415 391 5400 main
ebayley@keker.com | [vcard](#) | keker.com

From: Reetu Sinha <RSinha@susmangodfrey.com>
Sent: Friday, January 10, 2025 6:07 PM
To: [Alejandra Salinas <ASalinas@susmangodfrey.com>](mailto:ASalinas@susmangodfrey.com); [Edward A. Bayley <EBayley@keker.com>](mailto:EBayley@keker.com); openaicopyrightlitigation.lwteam@lw.com; [KVP-OAI <KVPOAI@keker.com>](mailto:KVPOAI@keker.com)
Cc: [Alter-TT \(Alter-AI-TT@simplelists.susmangodfrey.com\) <Alter-AI-TT@simplelists.susmangodfrey.com>](mailto:Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>); [OpenAICopyright <OpenAICopyright@mofo.com>](mailto:OpenAICopyright@mofo.com)
Subject: RE: AG/OAI: Source Code

[EXTERNAL]

Ed,

While we categorically maintain that these categories of source code are responsive to existing

requests, to preserve our rights, please find attached for service RFPs encompassing these categories of source code. We stand by the requests we made on our call yesterday, memorialized in Alejandra's email below.

Reetu Sinha (She/Her) | Attorney
p 713.655.5692
rsinha@susmangodfrey.com

From: Alejandra Salinas (via Alter-AI-TT list) <Alter-AI-TT@simplelists.susmangodfrey.com>
Sent: Friday, January 10, 2025 12:42 AM
To: Edward A. Bayley <EBayley@keker.com>; openaicopyrightlitigation.lwteam@lw.com; KVP-OAI <KVPOAI@keker.com>
Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>; OpenAICopyright <OpenAICopyright@mofo.com>
Subject: Re: AG/OAI: Source Code

EXTERNAL Email

Ed,

Thank you for our discussion today and for OAI's representation that it intends to produce all relevant source code history, which OAI anticipates will include the relevant source code about Books1 and Books2. As I said on the call, however, Plaintiffs remain concerned about the timeliness of OAI's response to the specific deficiencies I identified two weeks ago given the current discovery deadline. Please confirm by no later than Sunday whether OAI will produce the specific source code and information requested below.

Alejandra C. Salinas | Partner | Susman Godfrey LLP
o. 713.653.7802 | c. 956.337.0910

From: Edward A. Bayley <EBayley@keker.com>
Sent: Thursday, January 9, 2025 12:51 AM
To: Alejandra Salinas <ASalinas@susmangodfrey.com>; openaicopyrightlitigation.lwteam@lw.com <openaicopyrightlitigation.lwteam@lw.com>; KVP-OAI <KVPOAI@keker.com>
Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>; OpenAICopyright <OpenAICopyright@mofo.com>
Subject: RE: AG/OAI: Source Code

EXTERNAL Email

+MoFo (who had been left off this chain)
Alejandra,

Your position below makes no sense. What relief exactly do you plan to seek from the Court—i.e., what is the discovery request (or requests) for which you intend to file a motion to compel source code? OpenAI has already indicated that it is willing to produce relevant source code that is responsive to an RFP. In the absence of an RFP covering the source code requested, there cannot—*by definition*—be any “deficiency” in the production of that source code. Indeed, it appears that Plaintiffs are attempting to bypass the discovery process entirely, which is improper.

Moreover, your contention that Plaintiffs made a valid print request on December 18 is belied by the language of the Protective Order and the written record. To make a valid print request, the Protective Order expressly requires that a receiving Party *not only* save files in the PrintRequest folder *but also* that the Receiving party (i) “identify[] the files by file path, file name, and line numbers, and [ii] notify the producing Party of a request to print *those pages*.” Protective Order at sec. 10(k) (emphasis added). These explicit notice provisions exist for a reason—namely, to prevent this sort of gamesmanship. This is why OpenAI reached out on December 20 for confirmation of what line numbers of which files (if any) Plaintiffs wish to print, as required by the Protective Order, and you do not contest that Plaintiffs never responded to that email. Accordingly, Plaintiffs have to date still not complied with the requirements in the Protective Order. Nonetheless, although the Protective Order does not entitle a party to printouts of entire source code files just by saving them in the PrintRequest folder, in the interest of moving things forward, OpenAI will (based on your email) process your request to print the entire contents of all files saved in the folder, though not on the arbitrary timeline set forth in your email.

Regards,

-Ed

Edward A. Bayley

Keker, Van Nest & Peters LLP

633 Battery Street

San Francisco, CA 94111-1809

415 773 6672 direct | 415 391 5400 main

ebayley@keker.com | [vcard](#) | keker.com

From: Alejandra Salinas <ASalinas@susmangodfrey.com>

Sent: Wednesday, January 8, 2025 11:44 AM

To: Edward A. Bayley <EBayley@keker.com>; openaicopyrightlitigation.lwteam@lw.com; KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: Re: AG/OAI: Source Code

[EXTERNAL]

Ed,

We have repeatedly asked OpenAI to confirm whether it will remedy specific deficiencies Plaintiffs identified in the source code OpenAI provided for inspection and to meet and confer regarding these specific deficiencies. In response, OpenAI has repeatedly refused to confirm whether they will remedy these specific deficiencies and whether they will confer about these specific deficiencies. Your representation that “OpenAI will agree to make available for inspection relevant source code that is responsive to an RFP” is not responsive to these specific requests. Given OpenAI’s repeated refusal, Plaintiffs can only conclude that OpenAI is refusing to remedy and confer about these specific deficiencies, and we will plan to seek relief from the Court. Please be prepared to explain the basis for OpenAI’s refusal to engage on these specific deficiencies on tomorrow’s call.

As for the print request, your representation about what Plaintiffs requested is incorrect. On December 18, at the conclusion of our inspection, my colleague Reilly Stoler requested that OpenAI produce—as a print request—all the material that was saved to the print request folder. The discreet files in the PrintRequest folder are inherently identified by page and line number and comport with the print limitations set forth in the Protective Order. OpenAI’s refusal to process and produce our request is a violation of the Protective Order, which mandates that the requested print material be provided within 5 business days of the request. Please confirm that OpenAI will produce all of the requested printed materials by no later than January 10.

Alejandra C. Salinas | Susman Godfrey LLP
o. 713.653.7802 | c. 956.337.0910

From: Edward A. Bayley <EBayley@keker.com>

Date: Tuesday, January 7, 2025 at 1:01 AM

To: Alejandra Salinas <ASalinas@susmangodfrey.com>,

openaicopyrightlitigation.lwteam@lw.com

<openaicopyrightlitigation.lwteam@lw.com>, KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (<Alter-AI-TT@simplelists.susmangodfrey.com>) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: RE: AG/OAI: Source Code

EXTERNAL Email

Alejandra,

Regarding the “requests,” I have already responded to your email on this point. OpenAI will agree to make available for inspection relevant source code that is responsive to an RFP. We have agreed to meet-and-confer regarding Plaintiffs’ RFPs on January 9 and are preparing to address at that time any contentions that the existing RFPs identified in Ms. Lepic’s email cover additional source code.

With respect to the Google doc, although your expert uploaded some documents to the PrintRequest subfolder on the source code machine, Plaintiffs have yet to request printouts be

generated under the Protective Order. Andrew Perito emailed your colleague on December 20 asking for confirmation of whether Plaintiffs intended to make a print request—and, if so, to provide the page and line numbers requested for printing as required by the Protective Order—and we never received a response. Are Plaintiffs now requesting that the entire contents of all of the files in the PrintRequest folder be printed?

Thanks,

-Ed

Edward A. Bayley

Keker, Van Nest & Peters LLP
633 Battery Street
San Francisco, CA 94111-1809
415 773 6672 direct | 415 391 5400 main
ebayley@keker.com | [vcard](#) | keker.com

From: Alejandra Salinas <ASalinas@susmangodfrey.com>

Sent: Monday, January 6, 2025 9:48 AM

To: Edward A. Bayley <EBayley@keker.com>; openaicopyrightlitigation.lwteam@lw.com; KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: Re: AG/OAI: Source Code

[EXTERNAL]

Ed,

Since we have not heard anything further from OAI, please provide times to confer tomorrow or Wednesday as requested over a week ago.

Alejandra

Alejandra C. Salinas | Susman Godfrey LLP

o. 713.653.7802 | c. 956.337.0910

From: Alejandra Salinas <ASalinas@susmangodfrey.com>

Date: Thursday, January 2, 2025 at 4:32 PM

To: Edward A. Bayley <EBayley@keker.com>, openaicopyrightlitigation.lwteam@lw.com <openaicopyrightlitigation.lwteam@lw.com>, KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: Re: AG/OAI: Source Code

Ed,

As requested below, please confirm by no later than January 6 that OpenAI will provide all of the source code and information requested below for inspection. If not, please let us know your availability to confer on January 6 or 7.

Regarding the google doc, we will be able to provide additional information as soon as OpenAI provides the printed materials our expert requested from his inspection. Please let us know when you will be able to do so.

Alejandra

Alejandra C. Salinas | Susman Godfrey LLP
o. 713.653.7802 | c. 956.337.0910

From: Edward A. Bayley <EBayley@keker.com>

Date: Tuesday, December 31, 2024 at 9:36 PM

To: Alejandra Salinas <ASalinas@susmangodfrey.com>, openaicopyrightlitigation.lwteam@lw.com

<openaicopyrightlitigation.lwteam@lw.com>, KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (<Alter-AI-TT@simplelists.susmangodfrey.com>) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: RE: AG/OAI: Source Code

EXTERNAL Email

Alejandra,

OpenAI's agreements to produce source code (and other documents) are set forth in its objections and responses to Plaintiffs' discovery requests. Your colleague, Ms. Lepic, already reached out—on the same day as your email—about OpenAI's responses to discovery requests that Plaintiffs contend relate to source code, and we have already indicated to you that we are endeavoring to respond to those demands by next week. *See attached*. We will respond at the same time to your inquiry regarding "[REDACTED]" and "[REDACTED]," which were also mentioned in Ms. Lepic's email.

We cannot respond to your inquiry about the so-called "private Google doc" without more information of what you are talking about. Please provide the name of the document and/or where it was referenced in the source code so that we can investigate further.

Thanks,

-Ed

Edward A. Bayley

Keker, Van Nest & Peters LLP

633 Battery Street

San Francisco, CA 94111-1809

415 773 6672 direct | 415 391 5400 main

ebayley@keker.com | [vcard](#) | keker.com

From: Alejandra Salinas <ASalinas@susmangodfrey.com>

Sent: Thursday, December 26, 2024 2:10 PM

To: openaicopyrightlitigation.lwteam@lw.com; KVP-OAI <KVPOAI@keker.com>

Cc: Alter-TT (Alter-AI-TT@simplelists.susmangodfrey.com) <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: AG/OAI: Source Code

[EXTERNAL]

Counsel,

We write to follow up on several deficiencies and omissions our expert identified during our source code review. Please confirm by no later than January 6 that OpenAI will provide the source code and information requested below for inspection. If not, please let us know your availability to confer on January 6 or 7.

Source Code

- All source code related to LibGen, Books1, and Books2, including, but not limited to, the code that downloaded "libgen-1-dedup" and "libgen-2-dedup" data from the Libgen websites and mirrors, code that discovered the list of books to download, scraped the webpages, and filtered and processed the data;
- All source code related to "[REDACTED]", including, but not limited to, the code that discovered the list of books to download, crawler code, OCR code, and any filtering and processing code;
- All source code related to Common Crawl and ELibra, including, but not limited to, the code that discovered the list of books to download, crawler code, OCR code, and any filtering and processing code;
- All source code related to "[REDACTED]", including, but not limited to, the code that discovered the list of podcasts to download, crawler code, transcription code, and any filtering and processing code;
- All source code related to downloading YouTube videos, including, but not limited to, code on transcribing and any arrangement with Microsoft or Google;

- All source code related to GPTBot, including, but not limited to, how it makes web requests, discovers pages to crawl, and processes the downloaded data;
- Model architecture code for GPT-3, GPT-4, GPT-4o that clearly define all the training layers and functions;
- Model training code that iteratively optimizes the model parameters given a set of training data, including, but not limited to, the code that feeds training data into the model and the optimization functions used to train the model;
- Model hosting and inference code, including, but not limited to, when a user sends a text prompt on ChatGPT interface, the code that process the prompt, call the model, and returns the response. Include code on processing the user prompt and the response;
- Any content moderation code, including, but not limited to, any code that detects whether a response should be returned to the user in ChatGPT; and
- To the extent not covered above, any other script that includes the term “book”, “books”, “copyright”, and/or “license”.

Training Data

- Training data associated with "[REDACTED]" and "[REDACTED]." There are several references to these datasets in the source code we inspected, but did not see either of these datasets in the training data OpenAI provided.

Other Information

- There is a private Google doc referenced in the code that contains information on the source and copyrightability of training data. This document was inaccessible during our source code review and has not been produced.

Alejandra C. Salinas | Partner | Susman Godfrey LLP

o. 713.653.7802 | c. 956.337.0910

[Bio](#) | [vCard](#) | she/her

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/confirm/?u=SntCTGorhR74gzVq8zpmIA3u0mONDWvJ>